

Diplôme d'Etudes Approfondies de Ressources Phytogénétiques et Interactions biologiques

Ecole Doctorale Biologie Intégrative

Analyse de la diversité génétique d'une population restauratrice de fertilité utilisée pour la création de variétés hybrides de riz (*Oryza sativa* L.)

Filloux Denis

Chantereau Jacques
CIRAD-CA
Avenue Agropolis – TA 70 / 01
34398 Montpellier Cedex 5

Soutenu à Montpellier
Le 31 Aout 2001

Résumé Afin d'analyser la diversité génétique d'une population restauratrice de fertilité utilisée pour la création de variétés hybrides de riz (*Oryza sativa* L.), le polymorphisme de 22 loci microsatellites a été examiné sur un échantillon de 150 plantes issues de cette population. L'absence de déficit en hétérozygote ($F_{is}=0,001$) a confirmé le comportement panmictique de la population rendue artificiellement allogame par l'utilisation d'un système de stérilité mâle monogénique. La diversité génétique de la population s'est avérée relativement élevée avec un nombre moyen d'allèles détectés par locus égal à 4,32 et un taux d'hétérozygotie de 0,437. Environ la moitié des allèles a été représentée dans la population à une fréquence inférieure à 10 %. La population, comparée à la diversité globale des riz cultivés d'origine asiatique, a montré une forte similitude avec les riz *indica*, mais quelques allèles spécifiques au groupe des *japonica* ont été trouvés. Les variétés *indica* IR5, IR8, IR36 et Peta se sont montrées les plus proches génétiquement de la population. Les 7 lignées A avec lesquelles la population est croisée pour l'obtention d'hybrides, se sont avérées appartenir pleinement au groupe des riz *indica*. Les distances génétiques qui existent entre la population restauratrice et les lignées A, ont été estimées. De tels résultats devraient permettre, d'une part, d'ajuster la richesse allélique de la population pour la rendre plus représentative des riz *indica* et, d'autre part, de mieux comprendre, aux travers de dispositifs expérimentaux adéquats, la relation qui existe chez le riz entre l'effet d'hétérosis observé chez un hybride et la divergence génétique de ses parents.

Abstract In order to analyze genetic diversity of a fertility restorer population used for the creation of hybrid varieties of rice (*Oryza sativa* L.), the polymorphism of 22 microsatellite loci was examined in a sample of 150 individuals originating from this population. The absence of heterozygote deficit ($F_{is}=0.001$) confirmed that the population artificially made allogamous by the use of a monogenic male sterility system, followed the Hardy-Weinberg principle. The genetic diversity of the population was relatively high with an average number of alleles detected per locus equal to 4.32 and a heterozygote frequency of 0.437. About half of the alleles were represented in the population at a frequency lower than 10 %. The population, compared to the global diversity of Asia originated cultivated rices, showed a strong similarity with *indica* rices, but some specific alleles to the group of *japonica* were found. The *indica* varieties IR5, IR8, IR36 and Peta revealed the closest genetic linkage to the population. The 7 lines A with which the population is crossed for the hybrids obtaining, was shown to completely belong to the group of *indica* rices. The genetic distances which exist between the restorer population and the 7 lines A, were estimated. Such results should allow, on one hand, the allelic wealth of the population to be adjusted, thus making it more representative of the *indica* rices and, on the other hand, to better understand, through adequate experimental devices, the relation which exists in rice between the heterosis effect observed in a hybrid and the genetic differentiation of their parents.

Mots clés *Oryza sativa* L. - Diversité génétique - Population - Sélection récurrente - Variétés hybrides

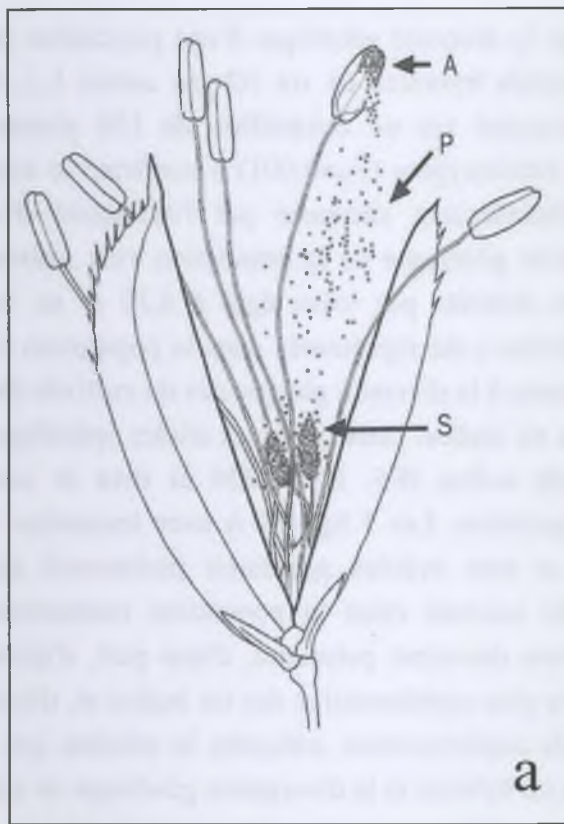


Figure 1 : Biologie florale et stérilité mâle génique chez le riz (A : anthère vide (v) ou remplie (r) de pollen ; P : pollen ; S : stigmat) :

a - schéma d'un épillet

b - panicule d'une plante mâle fertile

c - panicule d'une plante mâle stérile

Introduction

Le riz (*Oryza sativa* L.), première culture mondiale à vocation alimentaire, est consommé par plus de la moitié de la population mondiale et représente l'aliment de base de plus du tiers de la population. Principalement cultivé en Asie où 90 % de la production y est assurée, il représente la deuxième culture mondiale en terme de surface et de production. Le riz est une plante semi-aquatique qui est cultivé, pour 88 % des surfaces, en milieu irrigué ou inondé. Au cours des millénaires, son aire de culture, initialement cantonnée à l'Asie tropicale, s'est progressivement étendue à d'autres écosystèmes par limitation, notamment, des besoins en eau (culture pluviale) ou par tolérance accrue aux températures plus basses (zones d'altitude et zones à climat tempéré). Ainsi, de nos jours, c'est la culture pluviale qui s'est préférentiellement développée en Afrique et en Amérique latine et elle représente, dans ces continents, environ 60 % des surfaces cultivées en riz. Les zones à climat tempéré (pourtour méditerranéen, USA, cône sud latino-américain, Australie,...) voient également leur contribution à la production mondiale augmenter progressivement. Par ailleurs, des variétés adaptées à la culture en altitude ont récemment été développées à Madagascar (Jacquot *et al.*, 1997).

Plante hermaphrodite, le riz est considéré comme une plante autogame. Cependant, compte tenu de la biologie des épillets, l'autogamie n'est pas stricte et des allofécondations naturelles sont possibles (figure 1a). Plante annuelle, il présente un cycle de culture qui varie, selon les variétés, de 80 à 160 jours. Le mode de reproduction du riz cultivé se fait quasi exclusivement par graine et l'implantation de sa culture est faite soit par semis direct, soit par repiquage de plants obtenus en pépinière. Son génome diploïde ($2n = 24$) est relativement simple et présente peu de duplication. Le déterminisme génétique de nombreux caractères est connu. Le riz présente, par ailleurs, une diversité génétique importante. La collection internationale compte actuellement environ 80 000 accessions conservées à l'IRRI aux Philippines (Jackson, 1997). Des classifications morphologiques (Oka, 1958 ; Cheng *et al.*, 1984) puis moléculaire, basées sur le polymorphisme enzymatique (Glaszmann, 1987) visant à structurer la diversité génétique des riz, ont permis de distinguer 6 groupes dont 2 principaux : *indica* et *japonica*. Par la suite, d'autres marqueurs moléculaires plus polymorphes tels que les RFLP (Zhang *et al.*, 1992 ; Second et Ghesquière, 1995) ou les microsatellites (Patry-Bres, 2000) ont permis de confirmer et de préciser cette classification.

L'augmentation nécessaire de la production passe notamment par la création de variétés plus productives. Du fait de son autogamie, le riz est essentiellement cultivé sous forme de lignées pures. La méthode de sélection la plus pratiquée pour la création de variétés lignées est la sélection généalogique après croisement de deux géniteurs. D'autres méthodes complémentaires sont ponctuellement utilisées : la sélection en mélange (bulk), le retrocroisement, la mutagenèse induite, l'haplodiploïdisation par androgenèse (Jacquot *et al.*, 1997),... Par ailleurs, la sélection récurrente qui vise à mieux exploiter les ressources génétiques et à maintenir le potentiel de progrès génétique à un niveau suffisant, s'est développée ces 15 dernières années chez le riz (Borrero *et al.*, 1997). Initiée au Brésil en 1984 (Chatel et Guimaraes, 1995), elle est maintenant largement pratiquée dans tout le continent latino-américain et ponctuellement en Afrique et à Madagascar. Parallèlement, l'exploitation de la vigueur hybride, lié au phénomène d'hétérosis, décrite chez le riz par Jones (1926), a depuis longtemps été suggérée pour accroître les

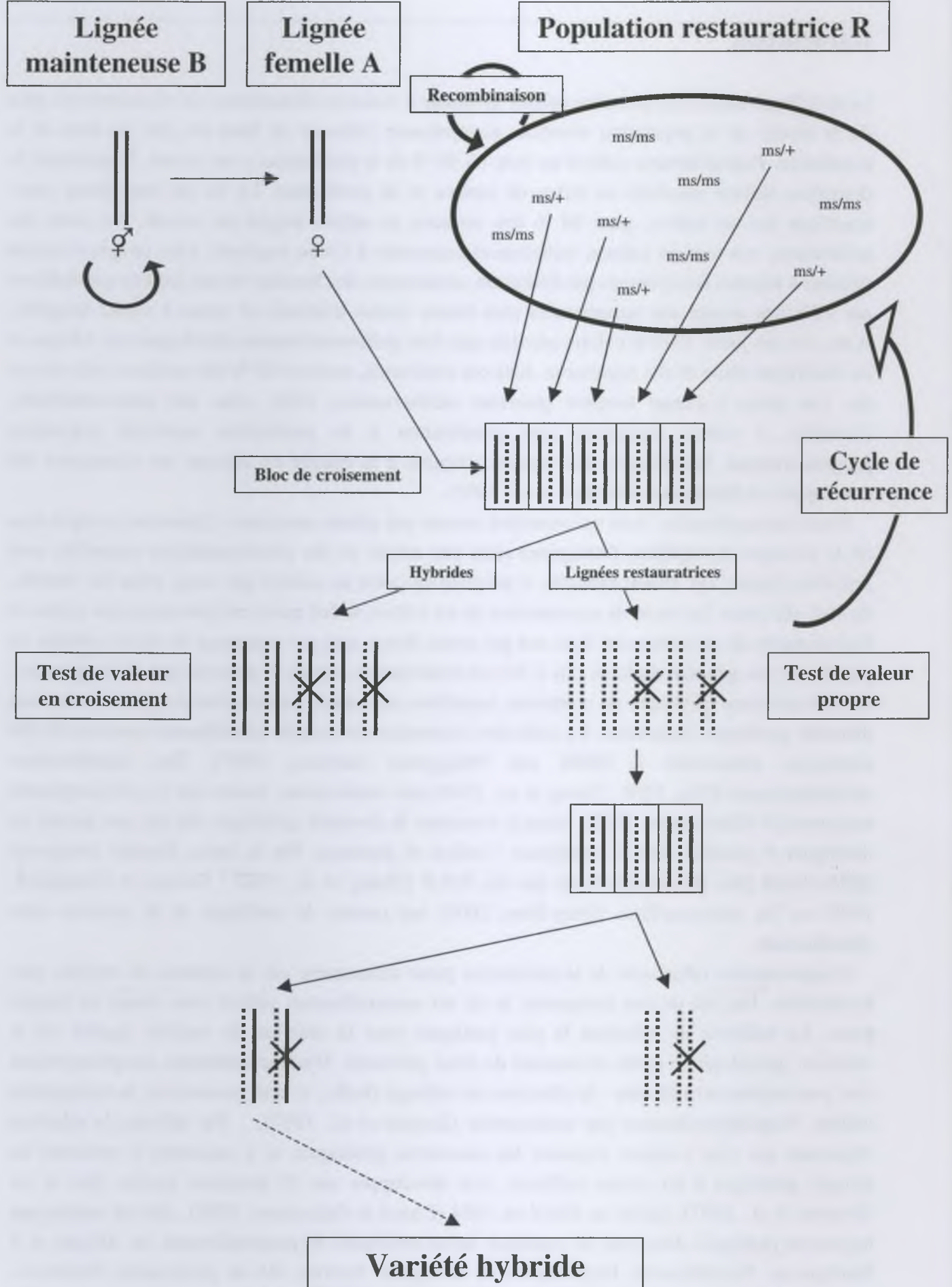


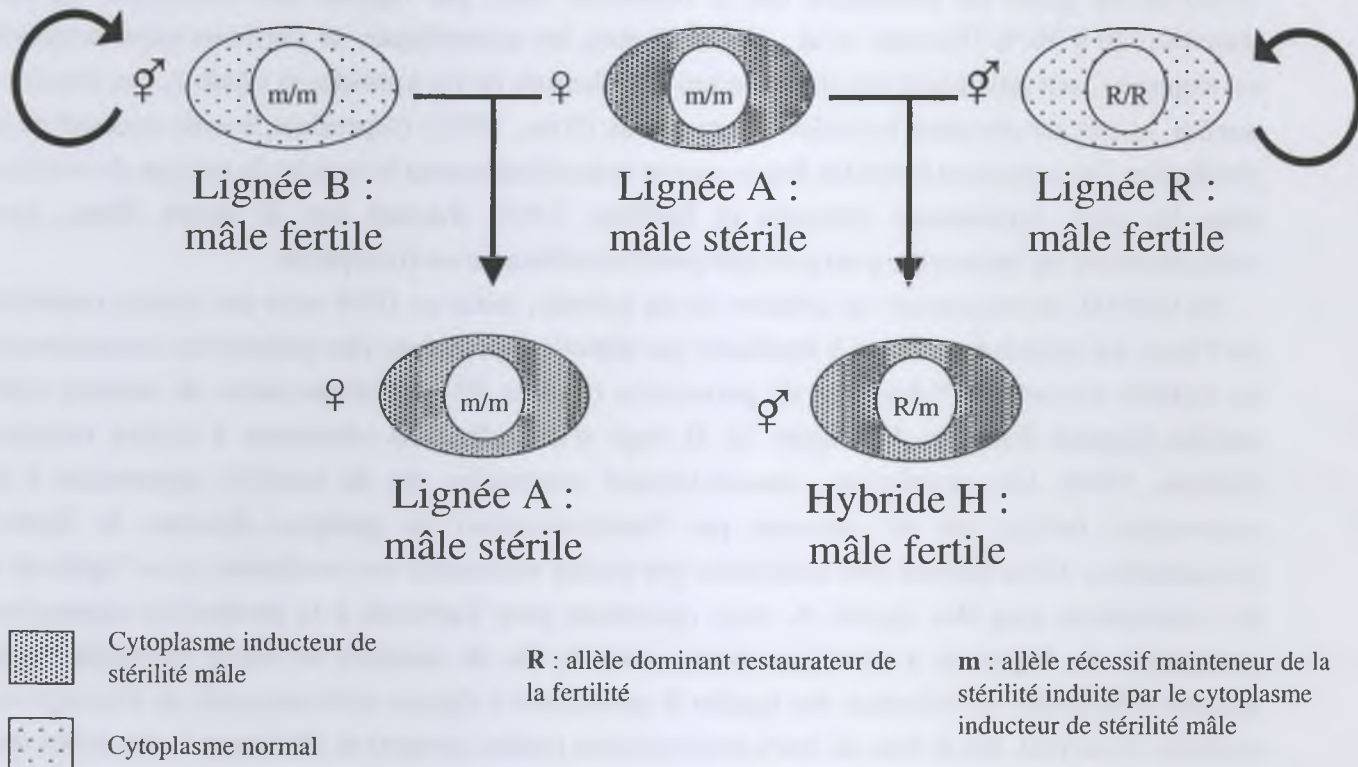
Figure 2 : Schéma simplifié de la création de variétés hybrides simples mettant en oeuvre la sélection récurrente à testeur constant pour l'obtention de parents mâles restaurateurs de fertilité (d'après Taillebois, comm. pers.)

rendements. Mais, c'est la découverte et la maîtrise de la stérilité mâle nucléo-cytoplasmique qui a permis, à partir de 1976, de développer en Chine les premières variétés hybrides élites (Lin et Yuan, 1980). Celles-ci représentent maintenant 55 % des surfaces en riz de ce pays (Zhang *et al.*, 1994) et les gains de rendement dus à l'hétérosis sont, par rapport aux meilleures lignées, d'environ 20 à 30 % (Virmani *et al.*, 1982). Depuis, les scientifiques de plusieurs pays, tempérés ou tropicaux, ont développé des programmes de sélection de riz hybrides et l'Inde a, ces dernières années, inscrit ses premiers hybrides commerciaux (Xiao, 1995). Cependant, le coût excessif de la production des semences hybrides freine encore considérablement le succès de ce type de matériel dans les pays occidentaux (Gravois et McNew, 1993) d'autant que le semis direct, gros consommateur de semences, y est pratiqué préférentiellement au riz repiqué.

Au CIRAD, le programme de création de riz hybride, initié en 1984 suite aux succès constatés en Chine, est raisonné de façon à améliorer par sélection récurrente des populations restauratrices de fertilité destinées à l'obtention de partenaires (lignées R) complémentaires de testeurs mâle stériles (lignées femelles A) (figure 2). Il s'agit d'une sélection récurrente à testeur constant (Gallais, 1989). Ces populations, essentiellement constituées par du matériel appartenant à la sous-espèce *indica*, ont été obtenues par l'intercroisement de quelques dizaines de lignées restauratrices. Elles doivent être améliorées par cycles successifs non seulement pour l'aptitude à la combinaison avec des lignées A, mais également pour l'aptitude à la production semencière (autrement dit, l'aptitude à l'allofécondation) dans le but de produire de façon économique les semences hybrides. La sélection des lignées R participant à chaque nouveau cycle de brassage est réalisée, d'une part, sur la base de leurs performances (valeur propre) et, d'autre part, sur celles des hybrides obtenus par leur croisement avec les lignées A (valeur en croisement). L'intercroisement entre individus au sein des populations restauratrices s'opère à l'aide d'une stérilité mâle génique provenant d'un mutant de la variété IR36 (Singh et Ikehashi, 1981) alors que la stérilisation des lignées A est assurée par l'utilisation d'une stérilité nucléo-cytoplasmique (cytoplasme WA) (Lin et Yuan, 1980) (figures 1 et 3).

Ce schéma de sélection étant très lourd, la mise au point d'une méthode plus simple d'identification des lignées R donnant des combinaisons hybrides performantes avec les lignées A, serait extrêmement utile. Pour parvenir à une telle capacité de prédiction, il est nécessaire de mieux connaître la relation qui existe entre l'effet d'hétérosis observé chez l'hybride et la divergence génétique de ses parents, ainsi que la part des différents effets génétiques supposés impliqués dans l'hétérosis (effets de dominance, de superdominance, d'épistasie,...). Il est soupçonné depuis longtemps que la vigueur hybride est liée à l'importance des différences génétiques existantes entre parents. Autrement dit, il existerait, chez l'hybride, une forte corrélation entre hétérosis et hétérozygotie. La confirmation d'une telle relation permettrait de prédire la valeur des hybrides à partir de la caractérisation moléculaire des parents. Le développement récent des marqueurs moléculaires a permis d'initier de nombreux travaux, principalement chez le maïs (Charcosset *et al.*, 1994), mais aussi chez le riz (Zhang *et al.*, 1996 ; Saghai Maroof *et al.*, 1997 ; Liu et Wu, 1998 ; Zhao *et al.*, 1999). Les résultats sont divers voire contradictoires, mais ils font néanmoins ressortir que les corrélations entre hétérosis et hétérozygotie sont d'autant plus marquées que l'on réalise les études avec des hybrides consanguins (croisements entre lignées apparentées). Ceci a été récemment vérifié chez le riz (Xiao *et al.*, 1996 ; Zhao *et al.*, 1999) avec des hybrides d'un même groupe sub-spécifique et de

Stérilité mâle nucléocytoplasmique utilisée pour la production des semences hybrides :



Stérilité mâle génique utilisée pour forcer les recombinaisons dans la population restauratrice :

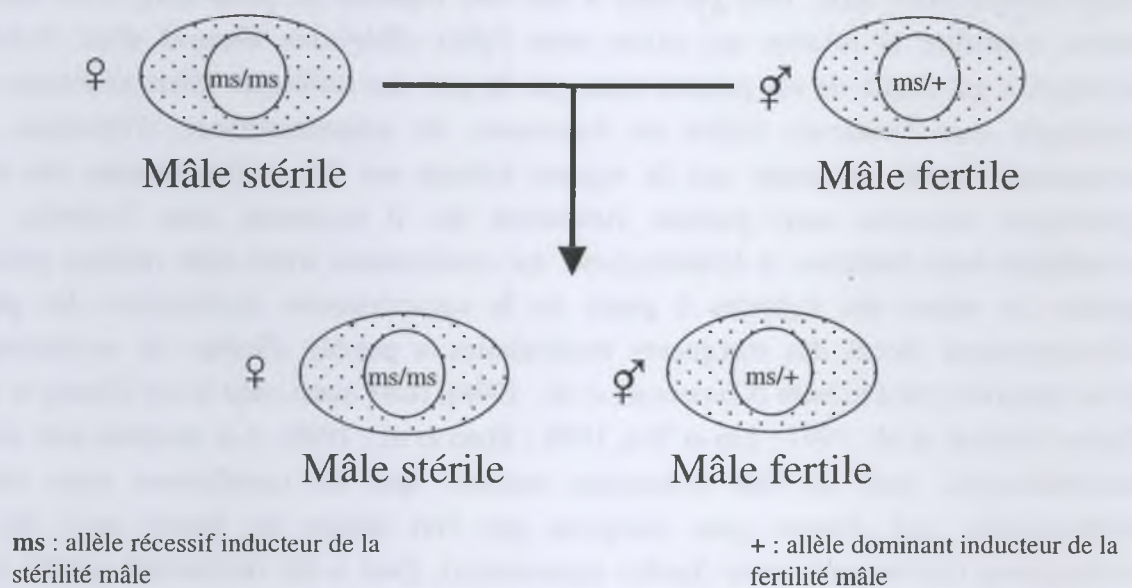


Figure 3 : Les deux types de stérilité mâle utilisées dans le schéma de création de variétés hybrides

groupes sub-spécifiques différents. Les auteurs montrent que l'hétérosis est maximum avec les hybrides *indica* x *japonica* mais, les corrélations entre distance moléculaire des parents et hétérosis sont, en revanche, faibles. Celles-ci ne sont élevées qu'avec les hybrides intra *japonica* et sont intermédiaires pour les croisements intra *indica*. Par ailleurs, pour ce type d'étude, les marqueurs de type RFLP, RAPD et, surtout, microsatellites, parce qu'ils couvrent mieux le génome et présentent un nombre d'allèles par locus plus important, semblent apporter des prédictions bien meilleures que les marqueurs isoenzymatiques (Charcosset et Gallais, 1998). La recherche de QTL (quantitative trait loci) présentant des effets d'hétérosis a été récemment mise en oeuvre sur le riz par Xiao *et al.* (1995). Ils révèlent que c'est la complémentation à des gènes dominants qui serait le facteur principal de l'hétérosis chez cette plante. Yu *et al.* (1997) montrent, en revanche, que l'épistasie joue un rôle majeur dans l'interprétation des bases génétiques de l'hétérosis.

Dans ce contexte, l'emploi de marqueurs moléculaires est envisagé pour aider à choisir les partenaires restaurateurs sélectionnés à chaque cycle de sélection récurrente pour leur aptitude à la combinaison. Cependant, avant d'engager une telle démarche, il est important, dans un tout premier temps, de mieux évaluer l'importance et la représentativité de la diversité génétique des populations restauratrices utilisées par rapport à la diversité globale des riz et par rapport à celle des testeurs. Dans cette étude, nous présentons les résultats de l'analyse de la diversité génétique d'une population restauratrice de fertilité utilisée par le programme de création de variétés hybrides de riz mené au Brésil par le CIRAD en se basant sur le polymorphisme de loci microsatellites. La diversité de la population est comparée, d'une part, à la diversité des lignées A disponibles et, d'autre part, à celle d'un échantillonnage représentatif des riz asiatiques cultivés.

Matériels et méthodes

Matériel végétal

Trois types de matériel ont été analysés :

- 150 plantes mâle fertiles (ms/+) choisies au hasard parmi environ 3000 plantes de la population restauratrice 00CP.

Pour constituer la population 00CP, 28 variétés et lignées ont été croisées chacune avec quelques plantes mâle stériles (ms/ms) prises au hasard dans la population CNA-IRAT 4 (Chatel et Guimaraes, 1995)(tableau 1). Les plantes (ms/+) issues de ces croisements ont été autofécondées et les graines récoltées ont été mélangées en quantité équivalente. Les graines récoltées sur les plantes mâle stériles issues de ce mélange constitue la population 00CP. A chaque cycle de brassage, une population est constituée de 3000 plantes environ. Les plantes sont cultivées en carré pour assurer un brassage uniforme. Elles sont aussi isolées de toute autre parcelle de riz à épiaison pour éviter toute contamination par le pollen. Seules les graines portées par les plantes mâle stériles sont récoltées pour reconduire la population. A l'équilibre, la population est constituée théoriquement de 50 % de plantes mâle fertile (ms/+) et de 50 % de plantes mâle stériles (ms/ms). Au moment de l'étude, la population était à son 5^{ème} cycle de brassage depuis sa création.

Tableau 1 : Liste des lignées fondatrices des populations 00CP et CNA-IRAT 4* et leurs participations relatives

Nom de la lignée	Origine génétique	Participation dans 00CP (%)	Participation dans CNA-IRAT4 (%)
1C.13.1.1.1	ZHEN SHAN 97A/METICA 1	1,79	-
1C.5.3.2.1	ZHEN SHAN 97A/METICA 1	1,79	-
CNA 3450	CICA7//CICA8/REMADJA	1,79	-
CNA 5682	-	1,79	-
CT 5747.12.9.3.7 = TOLIMA	COLOMBIA 1//P1274.6.8M.1.3M/P1274.6.8M.1.3M.1//P2060F4.2.5.2	1,79	-
IR70	IR19660.73.4/IR54/IR9828.36.3	1,79	-
IR 36	CR94-13/IR2042	1,79	-
ITA 212	BG90-2*4/TETEP	1,79	-
ELONI	ACORNI//KAPURI/IR454	1,79	-
BR.IRGA 408	IR8/IR12.178.2.3	1,79	-
C 1321.90R	-	1,79	-
CNA 5787	-	1,79	-
CNA 6816	-	1,79	-
CT 6163.8.9.1.2A	IR46/IRAT120//P1274.6.8M.1.3M.1	1,79	-
L306.11	Mutant d'ELONI	1,79	-
LA 110	H4/TAICHUNG NATIVE 1	1,79	-
METICA 1	F7 P738.137.3.1/P881.19.14.10//P738.137.3.1/P868B.24.5	1,79	-
SR.B.1.6	ZHEN SHAN 97A/METICA 1//CNA92.BM11.BM19p18	1,79	-
ST.B.14.2	ZHEN SHAN 97A/METICA 1//CNA4110	1,79	-
C 74	-	1,79	-
CT 6516.23.8.1	TOX1010.49.1/IRAT121//IRAT216	1,79	-
DIWANI	WASHABO/IR454.1.17.1.1	1,79	-
IRAT 216	COLOMBIA 1/M312A	1,79	-
IRI 378R	-	1,79	-
P 4134.F3.22.1.B	METICA 1//ELONI/S12.30	1,79	-
SQ.B.5.3	ZHEN SHAN 97A/METICA 1//CNA4120	1,79	-
IRAT 348	Précoce de MACHQUES/D52.37//ZENITH/NIRA//D85.42/4/CENTURY PATNA231/SLC...	1,79	-
SZ.B.9.1	ZHEN SHAN 97A/METICA 1//CNA4114	1,79	-
BG 90.2	IR 262/REMADJA	4,17	8,33
CNA 7	T 141/IR 665.1.1.75.3	4,17	8,33
CNA 3815	CICA 4/BG90.2//SML 5617	4,17	8,33
CNA 3848	IR 36/CICA 7//5461	4,17	8,33
CNA 3887	BG90.2/TETEP//4440	4,17	8,33
COLOMBIA 1	NAPAL/TAKAO IKU 18	4,17	8,33
ELONI	ACORNI//KAPURI/IR454	4,17	8,33
NANICÃO	Cultivar traditionnel brésilien	4,17	8,33
UPR 103.30.1.2	IR24/CAUVERY	4,17	8,33
IR 36ms	Mutant de IR 36 porteur du gène de stérilité mâle génique ms	12,50	25,00

* Chatel et Guimaraes (1995)

- 7 lignées A (tableau 2).

Ces lignées, obtenues par l'IRRI ou par le CIRAD, sont porteuses du cytoplasme WA.

- 85 variétés issues de la collection MiniGB (annexe).

La collection MiniGB est constituée de 270 variétés cultivées (Bonman *et al.*, 1990 ; Glaszmann *et al.*, 1995) en guise de référence de la diversité géographique, écotypique et enzymatique des riz cultivés d'origine asiatique *O. sativa*. Les 85 variétés ont été choisies dans chacun des 6 groupes identifiés par Glaszmann (1987) selon leur profil enzymatique. Cinquante-deux appartiennent au groupe I (*indica*), 19 aux groupes intermédiaires II (aus), III (riz flottants précoces), IV (riz flottants type Rayada), V (basmati), et 14 au groupe VI (*japonica*). Afin de mieux explorer la diversité du groupe *indica* auquel appartiennent a priori la population 00CP et les lignées A, ce groupe a été volontairement surreprésenté.

Pour chaque lignée A et chaque variété de la collection MiniGB, l'échantillon analysé est constitué par la plante issue d'une graine choisie au hasard dans le lot de semences disponibles.

Analyses moléculaires

Les données moléculaires concernant les 33 variétés des groupes II à VI ont été extraites des données collectées par Glaszmann *et al.* (1999) et par Luce *et al.* (2001). Pour le reste du matériel, les analyses ont été réalisées comme suit. L'ADN de chaque échantillon a été extrait à partir de feuilles prélevées sur les plantes en floraison et broyées dans de l'azote liquide selon la méthode décrite par Risterucci *et al.* (2000).

Pour étudier le polymorphisme des longueurs des sites microsatellites, 22 couples d'amorces ont été utilisés (tableau 3). Ces amorces ont été choisies pour la bonne répartition sur le génome des loci correspondants et pour le polymorphisme minimum que les loci présentent (Wu et Tanksley, 1993 ; Panaud *et al.*, 1996 ; Chen *et al.*, 1997 ; Temnykh *et al.*, 2000). L'amplification de l'ADN par PCR (Polymerase Chain Reaction) a été réalisée selon Chen *et al.* (1997) excepté que les amorces ont été préalablement marquées par phosphorylation en position 5' par du γ -[P³³]-ATP en utilisant une T4 polynucléotide kinase (Gibco BRL). Les produits de PCR ont été séparés par électrophorèse pendant 1 h 30 environ à 55 W sur gel dénaturant de polyacrylamide à 5 % contenant 7,5 % d'urée et baignant dans une solution tampon de TBE à 0,5%. Un marqueur de taille d'une gamme de 30 à 330 pb, présentant une incrémentation de 10 pb, préalablement marqué par du γ -[P³³]-ATP, a été utilisé pour estimer la taille de chaque locus microsatellite amplifié (ou allèle). Après migration, les gels ont été séchés à 80°C pendant 30 mn et exposés pendant 24 h à des films radiographiques (Kodak X-OMAT) pour révéler les bandes des marqueurs.

Analyses des données

Pour chaque locus microsatellite analysé, la diversité génétique a été estimée en calculant l'index de Nei (1978) non biaisé défini comme suit :

$$H_e = (2n / (2n - 1)) (1 - \sum_k X_i^2) \quad \text{avec : } X_i = \text{fréquence de l'allèle } i,$$

Tableau 2 : Liste et origine des lignées A étudiées

Nom de la lignée	Origine génétique	Obtenteur
ANT 24	IR 62829	IRRI
ANT 66	IR 68897 p	IRRI
ANT 67	IR 68897 t	IRRI
ANT 771	CNA-IRAT 4/P 3059.F4.25.3.1	CIRAD
ANT 777	ECIA 89.J9.9.4.1/IR 58025	CIRAD
ANT 779	PUSA BASMATI/IR 58025	CIRAD
ANT 781	IR 58025/CIRAD 403	CIRAD

Tableau 3 : Caractéristiques des marqueurs microsatellites utilisés (Wu et Tanksley, 1993 ; Panaud *et al.*, 1996 ; Chen *et al.*, 1997 ; Temnykh *et al.*, 2000)

Loci	Numéro de chromosome	Position chromosomique (cM)	Taille de l'allèle (pb) chez IR36	Motif chez IR36
RM1	1	19,3	113	(GA) ₂₆
RM11	7	50,7	140	(GA) ₁₇
RM13	5	28,8	141	(GA) ₁₆
RM17	12	103,4	184	(GA) ₂₁
RM18	7	100,7	157	(GA) ₄ AA(GA)(AG) ₁₆
RM19	12	17,4	226	(ATC) ₁₀
RM21	11	85,8	157	(GA) ₁₈
RM22	3	12,5	194	(GA) ₂₂
RM25	8	53,5	146	(GA) ₁₈
RM50	6	49,3	201	(GA) ₁₅
RM53	2	29,8	158	(GA) ₁₄
RM106	2	126,5	297	(GAA) ₅
RM148	3	225,5	129	(GT) ₁₂
RM149	8	113,3	253	(AT) ₁₀
RM164	5	75,4	246	(GT) ₁₆ TT(GT) ₄
RM167	11	36,5	128	(GA) ₁₆
RM171	10	68,0	328	(GATG) ₅
RM219	9	9,6	202	(GA) ₁₇
RM222	10	0,0	213	(GA) ₁₈
RM246	1	110,7	116	(GA) ₂₀
RM257	9	62,1	147	(GA) ₂₄
RM261	4	14,9	125	(GA) ₈

n = nombre d'individus caractérisés,

k = nombre d'allèles.

H_e est le taux d'hétérozygotie attendu dans une population suivant la loi de Hardy-Weinberg. Si l'on considère tous les loci, la diversité de la population est la moyenne arithmétique de la diversité H_e à chacun des loci.

Pour apprécier les écarts à la panmixie dans la population 00CP à chacun des loci polymorphes, l'indice de fixation de Wright a été calculé comme suit :

$$F_{is} = 1 - H_o / H_e \quad \text{avec : } H_o = \text{taux d'hétérozygotie observé.}$$

La signification de ce paramètre a été testé avec les tests proposés par le logiciel GENEPOP version 3.3 (Raymond et Rousset, 1995). Ainsi, pour les loci présentant moins de 4 allèles, les tests ont été réalisés selon Louis et Dempster (1987) et, dans les autres cas, le test utilisé a été celui de Guo et Thompson (1992). Par ailleurs, l'excès ou le déficit en hétérozygote ont été testés avec le test proposé par Rousset et Raymond (1995).

Pour estimer les distances génétiques qui existent entre les lignées A et la population 00CP, une matrice de distances basée sur la distance non biaisée de Nei (1978) a été construite. Pour un locus, celle-ci est définie comme suit :

$$D_{xy} = - \ln (G_{xy} / (G_x G_y)^{1/2}) \quad \text{où } G_x = (2 n_x \sum_k X_i^2 - 1) / (2 n_x - 1),$$
$$G_y = (2 n_y \sum_k Y_i^2 - 1) / (2 n_y - 1),$$
$$G_{xy} = \sum_k X_i Y_i,$$

avec : X_i et Y_i = fréquences de l'allèle i dans les population X et Y,

n_x et n_y = nombre d'individus caractérisés dans les populations X et Y,

k = nombre d'allèles.

Si l'on considère tous les loci, les G_x , G_y et G_{xy} sont les moyennes arithmétiques des valeurs à chaque locus. Une classification ascendante hiérarchique a été réalisée à partir de cette matrice par la méthode UPGMA (Sneath et Sokal, 1973). La construction du dendrogramme a été réalisée avec le logiciel DARWIN version 3.6 développé par Perrier *et al.* (1999).

Les analyses factorielles des correspondances (AFC)(Benzécri, 1973) visant à décrire les associations entre individus, ont été réalisées à partir d'une matrice des fréquences alléliques (tableau disjonctif complet croisant individus et allèles) selon le codage suivant : 0 pour l'absence de l'allèle, 1 pour la présence de l'allèle à l'état hétérozygote et 2 pour la présence de l'allèle à l'état homozygote (She *et al.*, 1987). Ces analyses ont été réalisées avec le logiciel GENETIX version 4.02 (Belkhir *et al.*, 2001).

Résultats

Comportement de la population 00CP

La population 00CP a été semée en décembre 1999 en Guadeloupe (16°7'N – 61°37'W). Les premières floraisons ont commencé le 61^{ème} jour après le semis et les dernières au 85^{ème} jour. Les floraisons se sont échelonnées régulièrement sur 35 jours environ et les plantes ont présenté un cycle variant de 96 à 120 jours. A titre de comparaison, le cycle des 7 lignées A a varié de 109 à 121 jours.

Tableau 4 : Fréquences alléliques et génotypiques observées pour 22 loci microsatellites dans la population restauratrice 00CP

Loci	Fréquences génotypiques (%)							Fréquences alléliques (%)
RM1	Allèles ^a	80	90	106	112	114	122	
	80	0,0	0,7	0,0	2,7	0,0	0,0	1,7
	90	----	3,3	0,0	27,3	2,7	0,0	18,7
	106	----	----	0,0	1,3	0,0	0,0	0,7
	112	----	----	----	46,0	14,0	2,0	69,7
	114	----	----	----	----	0,0	0,0	8,3
	122	----	----	----	----	----	0,0	1,0
RM11	Allèles	125	131	141	145			
	125	0,0	2,7	0,0	0,0			1,3
	131	----	26,7	2,0	49,3			53,7
	141	----	----	0,0	2,0			2,0
	145	----	----	----	17,3			43,0
RM13	Allèles	129	141	143	145	155	159	
	129	0,0	2,7	4,7	0,0	17,3	0,0	12,3
	141	----	0,0	8,0	0,0	10,7	0,7	11,0
	143	----	----	6,0	0,7	26,0	0,7	26,0
	145	----	----	----	0,0	1,3	0,0	1,0
	155	----	----	----	----	21,3	0,0	49,0
	159	----	----	----	----	----	0,0	0,7
RM17	Allèles	160	186					
	160	2,7	31,3					18,3
	186	----	66,0					81,7
RM18	Allèles	150	156	158	160			
	150	0,7	4,0	22,7	1,3			14,7
	156	----	2,0	12,0	0,0			10,0
	158	----	----	56,0	1,3			74,0
	160	----	----	----	0,0			1,3
RM19	Allèles	214	217	244	250			
	214	0,0	0,0	6,7	0,0			3,3
	217	----	2,7	29,3	0,7			17,7
	244	----	----	60,0	0,7			78,3
RM21	Allèles	131	141	149	167	169		
	131	90,0	2,0	3,3	3,3	1,3		95,0
	141	----	0,0	0,0	0,0	0,0		1,0
	149	----	----	0,0	0,0	0,0		1,7
	167	----	----	----	0,0	0,0		1,7
	169	----	----	----	----	0,0		0,7
RM22	Allèles	192	194	196	198			
	192	0,0	0,0	1,3	0,0			0,7
	194	----	13,3	44,0	2,0			36,3
	196	----	----	34,7	4,7			59,7
	198	----	----	----	0,0			3,3
RM25	Allèles	146						
	146	100,0						100,0
RM50	Allèles	182	198	202				
	182	6,0	27,3	13,3				26,3
	198	----	24,0	23,3				49,3
	202	----	----	6,0				24,3

Tableau 4 : Suite

Loci	Fréquences génotypiques (%)								Fréquences alléliques (%)
RM246	Allèles	96	113	117					
	96	78,0	2,0	18,7					88,3
	113	----	0,0	0,0					1,0
	117	----	----	1,3					10,7
RM257	Allèles	143	144	146	148	174	182	186	
	143	6,7	0,0	0,7	37,3	0,7	0,0	0,7	26,3
	144	----	0,0	0,0	0,7	0,0	0,0	0,7	0,7
	146	----	----	0,0	0,0	0,0	0,0	0,0	0,3
	148	----	----	----	48,7	2,0	0,7	0,7	69,3
	174	----	----	----	----	0,0	0,7	0,0	1,7
	182	----	----	----	----	----	0,0	0,0	0,7
	186	----	----	----	----	----	----	0,0	1,0
RM261	Allèles	122	123	124	125	126	131		
	122	1,3	0,7	0,0	0,0	0,0	2,0		2,7
	123	----	6,0	4,0	0,0	1,3	23,3		20,7
	124	----	----	0,0	0,0	0,0	9,3		6,7
	125	----	----	----	0,0	0,0	0,7		0,3
	126	----	----	----	----	0,0	2,0		1,7
	131	----	----	----	----	----	49,3		68,0

^a Les allèles sont exprimés par le poids moléculaire en paire de bases (pb) du produit de la PCR

A la floraison, les plantes mâle stériles se sont aisément distinguées des plantes mâle fertiles par leurs anthères fines et blanches rendues ainsi par l'absence de grain de pollen (figure 1c). A maturité, le taux de fertilité paniculaire s'est élevé à un niveau appréciable d'environ 35 %.

Par ailleurs, sur 402 plantes observées, le taux de plantes mâle stériles s'est élevé à 47,0 %. Un test de χ^2 ($\chi^2 = 1,43$; ddl = 1) ne montre pas d'écart significatif ($P > 0,20$) avec la proportion théorique attendue équivalente à 50 %. La proportion en plantes mâle stériles et mâle fertiles est statistiquement conforme à ce qui était attendu.

Diversité génétique, fréquences alléliques et génotypiques de la population 00CP

Les fréquences alléliques et génotypiques de la population 00CP pour les 22 loci microsatellites sont présentées dans le tableau 4. Au total, 95 allèles ont été trouvés dans les 150 plantes échantillonnées dans la population 00CP. La plupart des loci se sont révélés polymorphes dans une gamme de 1 à 7 allèles par locus. Si l'on considère tous les loci, chaque plante analysée a présenté un génotype distinct. Le nombre d'allèles moyen par locus, égal à 4,32, s'est trouvé relativement élevé (tableau 5). Deux loci (RM25 et RM167) n'ont montré aucun polymorphisme alors qu'ils ont révélé un polymorphisme important dans la collection de référence de 85 variétés d'*O. sativa* analysée. A l'inverse, les 2 loci les plus polymorphes dans la population ont été RM222 et RM257, et faisaient aussi partie des loci les plus polymorphes si l'on considère l'ensemble des échantillons analysés.

Par ailleurs, une proportion importante d'allèles peu fréquents a caractérisé la population (figure 4). En effet, 46 % des allèles ont été représentés à une fréquence inférieure à 10 % et seulement 17 % des allèles ont été représentés à une fréquence supérieure à 50 %. Cette forte proportion en allèles peu fréquents a indiqué que la population 00CP disposait d'une diversité importante où peu d'allèles ont été fixés. Le taux d'hétérozygotie moyen H_e de la population, égal à 0,437, l'a confirmé.

Ecart à la panmixie pour la population 00CP

L'incorporation du gène de stérilité mâle nucléaire dans la population 00CP permet de contraindre les croisements entre plantes et doit permettre, en principe, le brassage aléatoire des gamètes. Pour vérifier si les intercroisements se sont bien produits de façon aléatoire, l'écart à la panmixie selon le principe de Hardy-Weinberg a été estimé et testé (tableau 6). Les écarts éventuels ont été mesurés par le calcul de l'indice de fixation de Wright F_{is} . D'une façon globale, la population, avec un indice de fixation quasiment nul ($F_{is} = 0,001$) n'a montré ni d'excès ni de déficit en hétérozygotes, suggérant que la conduite de la population, lors du cycle de brassage précédent, s'est opéré sans biais significatif. Cependant, certains loci ont exprimé ponctuellement une déviation par rapport à la panmixie. C'est le cas notamment des loci RM13 et RM222 qui ont montré un excès très significatif en hétérozygotes. Le locus RM261 a, quand à lui, montré un déficit en hétérozygotes très significatif.

Tableau 5 : Variation allélique et diversité génétique des loci microsatellites pour les différents types de matériels analysés

Loci	Population restauratrice 00CP		Lignées A		<i>Indica</i> (I)		<i>Japonica</i> (VI)		Groupes intermédiaires ^a		Total	
	(n ^b = 150)		(n = 7)		(n = 52)		(n = 14)		(n = 19)		(n = 242)	
	Nb allèles	He	Nb allèles	He	Nb allèles	He	Nb allèles	He	Nb allèles	He	Nb allèles	He
RM1	6	0,474	4	0,659	15	0,855	4	0,730	5	0,791	18	0,735
RM11	4	0,528	3	0,484	7	0,763	5	0,794	4	0,541	11	0,772
RM13	6	0,667	3	0,615	9	0,780	5	0,646	8	0,791	15	0,806
RM17	2	0,300	2	0,143	7	0,675	1	0,000	5	0,717	9	0,547
RM18	4	0,422	3	0,484	8	0,742	4	0,550	7	0,797	10	0,644
RM19	4	0,355	3	0,703	9	0,799	2	0,423	3	0,432	11	0,643
RM148	3	0,498	2	0,264	2	0,493	2	0,476	3	0,432	4	0,507
RM149	6	0,646	3	0,615	9	0,853	8	0,857	4	0,575	14	0,739
RM164	6	0,688	2	0,528	13	0,890	6	0,709	7	0,785	20	0,835
RM167	1	0,000	1	0,000	8	0,484	9	0,878	9	0,757	14	0,280
Sous-total	42		26		87		46		55		126	
Sous-moyenne	4,20	0,458	2,60	0,449	8,70	0,733	4,60	0,606	5,50	0,662	12,60	0,651
RM21	5	0,097	3	0,659	15	0,885	- ^c	-	-	-	17	0,410
RM22	4	0,513	3	0,703	4	0,527	-	-	-	-	4	0,551
RM25	1	0,000	1	0,000	10	0,840	-	-	-	-	10	0,313
RM50	3	0,630	1	0,000	11	0,778	-	-	-	-	11	0,693
RM53	6	0,686	4	0,659	8	0,792	-	-	-	-	8	0,766
RM106	2	0,480	2	0,264	4	0,604	-	-	-	-	4	0,509
RM171	3	0,649	2	0,264	5	0,315	-	-	-	-	5	0,669
RM219	6	0,302	4	0,747	15	0,911	-	-	-	-	18	0,580
RM222	7	0,527	4	0,659	13	0,871	-	-	-	-	13	0,659
RM246	3	0,209	3	0,703	7	0,725	-	-	-	-	7	0,418
RM257	7	0,451	4	0,659	9	0,811	-	-	-	-	12	0,614
RM261	6	0,491	3	0,484	5	0,583	-	-	-	-	6	0,662
Total	95		60		193						241	
Moyenne	4,32	0,437	2,73	0,468	8,77	0,726					10,95	0,607

^a Les groupes intermédiaires regroupent les groupes II (aus), III (riz flottants précoces), IV (riz flottants type Rayada) et V (basmati)

^b n = effectif

^c - = données manquantes

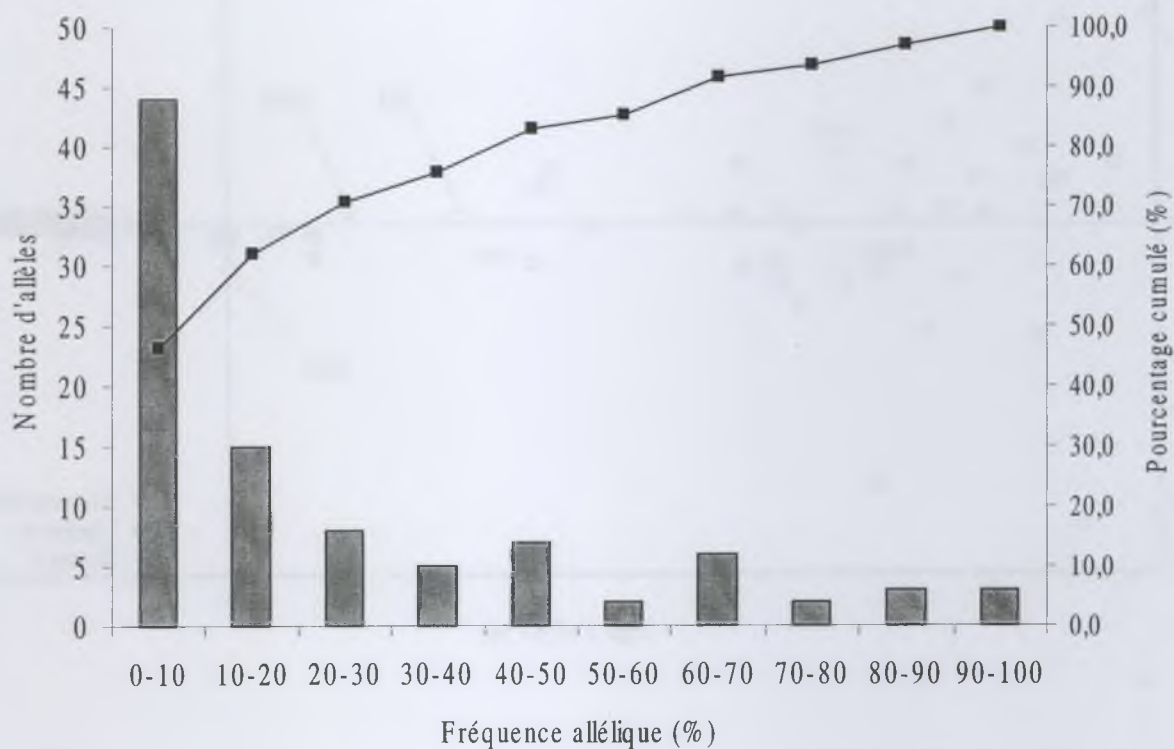


Figure 4 : Distribution des fréquences alléliques pour 22 loci microsatellites étudiés dans la population restauratrice 00CP

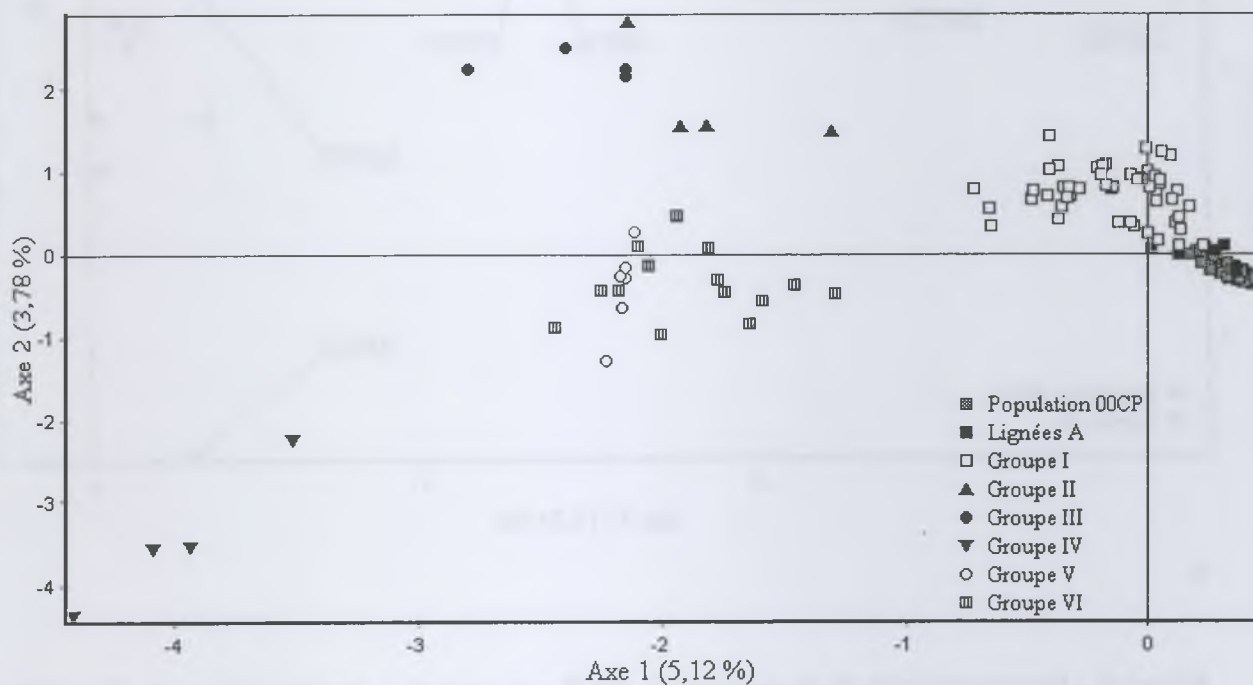
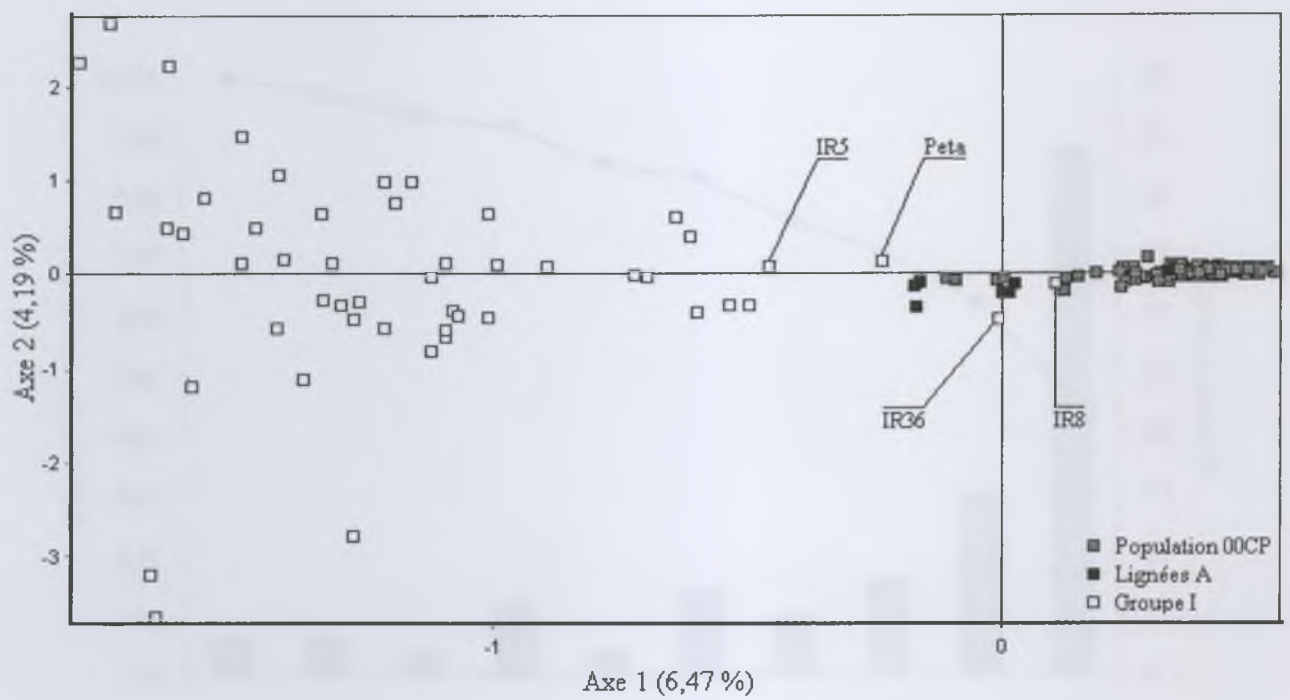
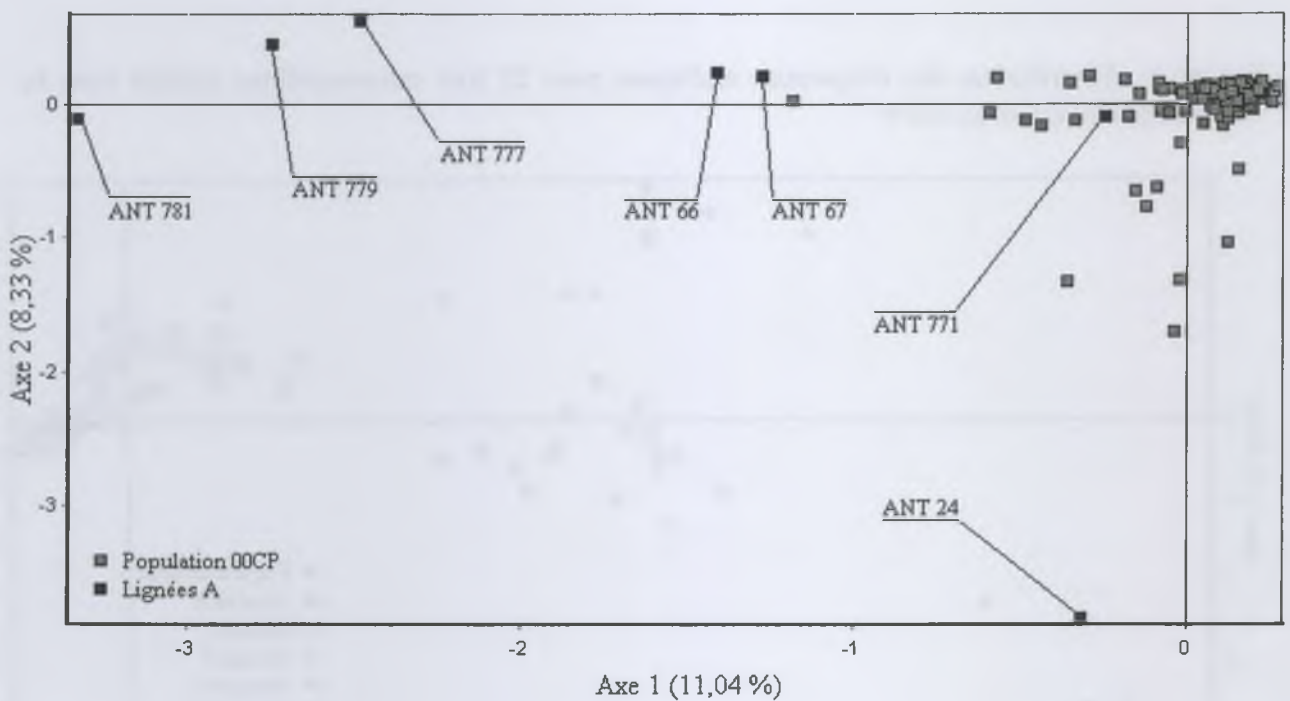


Figure 5 : Positionnement de la population 00CP et des 7 lignées A par rapport à un échantillon de 85 variétés représentatives de la diversité des riz *O. sativa*, révélé par analyse factorielle des correspondances basée sur les fréquences alléliques de 10 loci microsatellites. Les groupes sont définis selon la classification enzymatique établie par Glaszmann (1987)



a



b

Figure 6 : Positionnement de la population 00CP par rapport à un échantillon de 52 variétés *indica* (groupe I) et à 7 lignées A (a), et par rapport à ces mêmes lignées A (b), révélé par analyse factorielle des correspondances basée sur les fréquences alléliques de 22 loci microsatellites

Comparaison de la population 00CP et des lignées A par rapport à la collection de référence *O. sativa*

La comparaison de la diversité génétique de la population 00CP et des 7 lignées A avec celle de la collection de référence de 85 variétés d'*O. sativa* a été réalisée pour 10 loci microsatellites. Les résultats sont présentés dans le tableau 5. Pour les 10 loci étudiés, il a été recensé dans la population 00CP le tiers (42 exactement) des 126 allèles trouvés dans l'ensemble du matériel végétal étudié. La population 00CP a présenté 2 allèles spécifiques (RM13-159 pb et RM164-300 pb) que l'on n'a pas retrouvé dans la collection de référence ce qui est peu compte tenu du nombre d'allèles total trouvés et du nombre relativement faible d'échantillons analysés. Si la fréquence du premier allèle s'est montré très faible dans la population 00CP (0,7 %), le second allèle a été, en revanche, plus fréquent (12 %). Les lignées A n'ont pas présenté d'allèle spécifique. Parmi les 126 allèles, 71 (56,3 %) se sont révélés spécifiques de chacun des 6 groupes. Ainsi, 47 (37,3 %) étaient spécifiques des 52 variétés du groupe *indica* (groupe I), 11 (8,7 %) étaient spécifiques des 14 variétés du groupe *japonica* (groupe VI) et 13 (18,3 %) ont été spécifiques des 19 variétés des groupes intermédiaires (groupe II, III, IV et V). Si l'on ne considère que les allèles spécifiques de chacun des groupes, 75,0 % et 15,0 % des allèles de la population 00CP proviennent, respectivement, du groupe I et VI. Aucun allèle de la population 00CP ne provient des groupes intermédiaires. En ce qui concerne les lignées A, tous les allèles proviennent du groupe I.

Par ailleurs, une AFC a été réalisée pour établir les associations entre les individus, permettre une structuration de la diversité génétique globale et, surtout, situer la population 00CP et les 7 lignées A par rapport à la collection de référence (figure 5). La structure établie par Glaszmann (1987) selon les données enzymatiques est partiellement reconnaissable. Ainsi, parmi les 4 groupes détectables, le groupe I (*indica*) et le groupe IV s'opposent tout particulièrement. Ceux-ci forment l'axe 1 qui explique 5,12 % de la variation totale. Les 2 autres groupes formés respectivement par les groupes II et III et les groupes V et VI sont, quant à eux, plus difficiles à distinguer. La population 00CP est fortement marquée par son appartenance au groupe des variétés *indica*, mais elle ne se confond pas totalement avec ce groupe. Cette distinction par rapport aux groupes des variétés *indica* trouve plus son origine dans l'association d'allèles particuliers que dans la présence d'allèles spécifiques à la population 00CP. Les lignées A, quant à elles, se situent en position intermédiaire entre les variétés *indica* et la population 00CP.

Comparaison de la population 00CP par rapport aux variétés *indica* et aux lignées A

La comparaison de la diversité génétique de la population 00CP avec celles des 7 lignées A et l'échantillon de 52 variétés *indica* a été réalisée de façon plus approfondie pour 22 loci microsatellites. Les résultats sont présentés dans le tableau 5. Sur un total de 95 allèles recensés dans la population 00CP, 16 (16,8 %) ne se retrouvent pas dans l'échantillon de 52 variétés *indica*. Trois des 60 allèles des lignées A ne se retrouvent pas dans la population 00CP, mais, en revanche, se retrouvent tous dans l'échantillon des variétés *indica*. Que les analyses soient faites sur 10 ou 22 loci microsatellites, les résultats en terme de diversité (index de diversité, nombre

Tableau 6 : Ecart à la panmixie observé dans la population restauratrice 00CP

Loci	Nb d'allèles	Ho	He	Fis	Ecart Hardy-Weinberg ^a		
					Global	Déficit	Excès
RM1	6	0,507	0,474	-0,069	ns	ns	ns
RM11	4	0,560	0,528	-0,060	ns	ns	ns
RM13	6	0,727	0,667	-0,089	ns	ns	**
RM17	2	0,313	0,300	-0,043	ns	ns	ns
RM18	4	0,413	0,422	0,021	ns	ns	ns
RM19	4	0,373	0,355	-0,051	ns	ns	ns
RM21	5	0,100	0,097	-0,030	ns	ns	ns
RM22	4	0,520	0,513	-0,015	ns	ns	ns
RM25	1	-	-	-	-	-	-
RM50	3	0,640	0,630	-0,016	ns	ns	ns
RM53	6	0,733	0,686	-0,069	ns	ns	ns
RM106	2	0,420	0,480	0,125	ns	*	ns
RM148	3	0,507	0,498	-0,017	ns	ns	ns
RM149	6	0,520	0,646	0,195	*	*	ns
RM164	6	0,613	0,688	0,109	ns	*	ns
RM167	1	-	-	-	-	-	-
RM171	3	0,680	0,649	-0,048	ns	ns	ns
RM219	6	0,293	0,302	0,030	ns	ns	ns
RM222	7	0,600	0,527	-0,139	ns	ns	**
RM246	3	0,207	0,209	0,011	ns	ns	ns
RM257	7	0,447	0,451	0,010	*	ns	ns
RM261	6	0,433	0,491	0,118	ns	**	ns
Total	95	0,437	0,437	0,001	ns	ns	ns

^a *** $P < 0,01$; * $P < 0,05$; ns $P > 0,05$

moyen d'allèles par locus,...) sont similaires et n'apportent que peu d'informations supplémentaires. La population 00CP présente une diversité, en terme de nombre d'allèles et d'index de diversité, environ 2 fois moins grande que celle de l'échantillon des variétés *indica*.

La figure 6a représente une AFC réalisée pour établir les relations entre la population 00CP, les 7 lignées A et l'échantillon de variétés *indica*. Comme la première AFC l'a déjà montré (figure 5), la population 00CP se distingue de la plupart des variétés *indica*. Cette distinction est révélée par l'axe 1 qui explique 6,47 % de la variation totale. La aussi, cette distinction par rapport au groupe des variétés *indica* trouve plus son origine dans l'association d'allèles particuliers que dans la présence d'allèles spécifiques à la population 00CP. Seules quelques variétés présentent une forte similitude avec la population 00CP. Ainsi, les 4 variétés *indica* les plus proches sont IR36, IR 5, IR8 et Peta. Une telle proximité n'a rien d'étonnant. D'une part, ces variétés sont plus ou moins consanguines et Peta se retrouve à différent degré dans le pedigree des 3 premières variétés. D'autre part, IR36 a directement contribué, de façon relativement forte, à la constitution de la population 00CP et IR 8, elle-même parente de certaines lignées fondatrices de la population 00CP, y a contribué aussi de façon indirecte (tableau 1). Quant aux lignées A, l'AFC (figure 6a) a confirmé leur position intermédiaire entre les variétés *indica* et la population 00CP, et leur similitude relative avec les variétés IR36, IR 5, IR8 et Peta. Cependant, le pedigree mal connu des lignées A n'a pas permis de confirmer cette similitude.

La figure 6b représente une AFC réalisée pour établir les relations entre la population 00CP et les 7 lignées A. Pour compléter cette analyse, une classification ascendante hiérarchique a été réalisée par la méthode UPGMA (figure 7) à partir de la matrice des distances non biaisées de Nei (tableau 7). Dans les deux types d'analyses, la proximité génétique des lignées ANT 66, ANT 67 et ANT 771 avec la population 00CP a été montrée. Dans le cas de ANT 771, ce résultat s'explique bien par l'analyse de son pedigree (tableau 2). En effet, celle-ci est une lignée sélectionnée à partir de la population CNA-IRAT 4, elle-même composant à 50 % la population 00CP. A l'inverse, la lignée ANT 24 a présenté la plus grande distance avec la population 00CP. Par ailleurs, ANT 66 et ANT 67 sont très proches. Sur 22 loci, ces deux lignées A n'ont différé que pour 4 loci. Leur proximité est aussi confirmée par leur pedigree car elles sont toutes deux issues du même croisement. Une certaine proximité existe aussi entre ANT 777, ANT 779 et ANT 781 du fait de leur parent commun (IR58025).

Discussion

Allogamie et intercroisements dans la population 00CP

Le système de stérilité mâle génique utilisé au sein de la population 00CP pour contraindre les intercroisements semble efficace malgré la nature initialement autogame du riz. L'observation de cette population au cours d'un cycle de culture nous a permis de constater une proportion équivalente en plantes mâle stériles et mâle fertiles, conformément à la théorie. Le taux de fertilité paniculaire des plantes mâle stériles, égal à environ 35 %, est considéré comme suffisant (Neves *et al.*, 1994) pour permettre un brassage correct des plantes et une production suffisante de semences pour reconduire la population. Ce bon taux de fertilité paniculaire chez les plantes mâle

Tableau 7 : Matrice des distances génétiques non biaisées de Nei (1978) entre 7 lignées A et la population restauratrice 00CP

Population ou lignée A	00CP	ANT 24	ANT 66	ANT 67	ANT 771	ANT 777	ANT 779
ANT 24	0,709						
ANT 66	0,440	0,788					
ANT 67	0,439	0,693	0,201				
ANT 771	0,485	1,145	0,693	0,788			
ANT 777	0,681	1,705	0,693	0,693	0,788		
ANT 779	0,639	1,145	0,452	0,452	0,606	0,383	
ANT 781	0,921	1,196	0,816	0,816	0,716	0,465	0,465

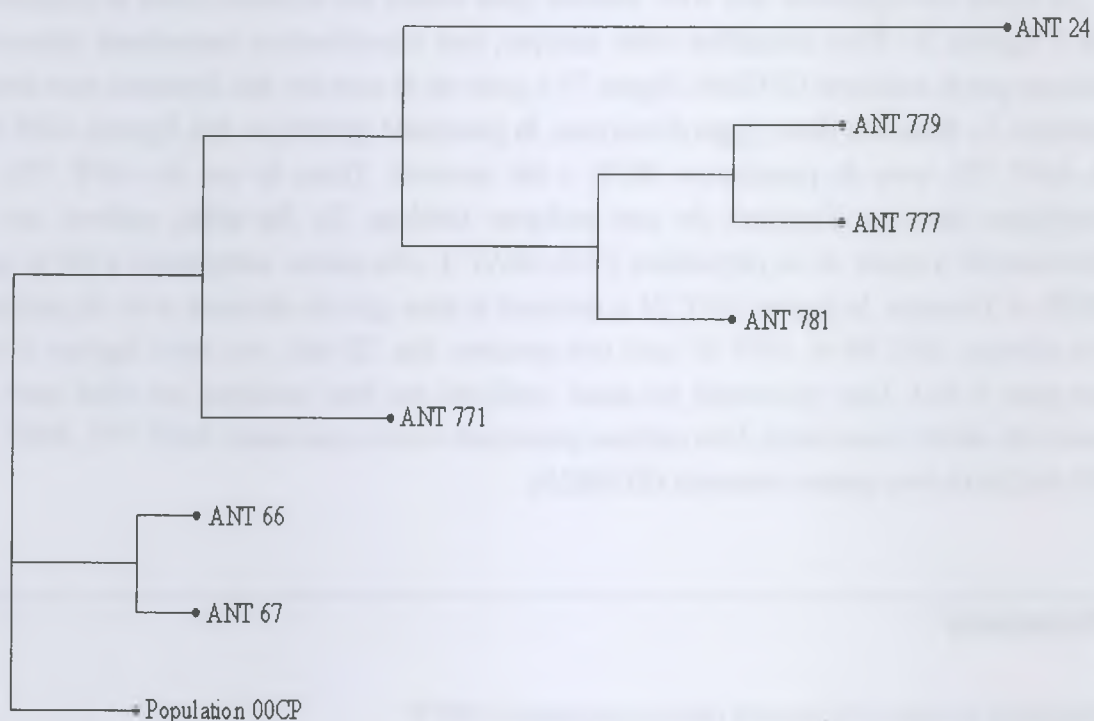


Figure 7 : Dendrogramme représentant la relation génétique entre la population 00CP et 7 lignées A, établi par méthode UPGMA à partir de la matrice des distances non biaisées de Nei (1978)

stériles s'explique en grande partie par la bonne exsertion des anthères associée à une production importante de pollen chez les plantes mâle fertiles, d'une part, et par la bonne exsertion des stigmates chez les plantes mâle stériles, d'autre part. Afin d'améliorer l'aptitude à l'allofécondation, l'introggression de caractères d'allogamie tels que les longs stigmates depuis *O. longistaminata* vers *O. sativa* avait été envisagée, à l'origine des programmes de sélection récurrente du riz, puis réussie par Taillebois (1983). Cependant, les caractères d'allogamie se sont rapidement avérés facilement sélectionnables au sein du matériel d'origine *indica* et le taux d'allofécondation s'est avéré suffisant (supérieur à 30 %) sans qu'une quelconque introgression n'ait été finalement nécessaire (Neves *et al.*, 1994).

Par ailleurs, l'absence d'écart global à la panmixie selon le principe de Hardy-Weinberg a été constatée malgré la présence ponctuelle de trois loci (RM13, RM222 et RM261) exprimant une déviation très significative par rapport à la panmixie, soit dans le sens d'un excès en hétérozygotes, soit dans le sens d'un déficit en hétérozygotes. Depuis la création de la population 00CP, une sélection volontaire, mais limitée, aura été pratiquée, au cours des différents cycles de culture, pour éliminer les plantes aberrantes pour leur architecture, la durée de leur cycle ou la morphologie du grain (format, aristation). En outre, ce mode de conduite a favorisé les plantes les plus prolifiques et les plus allogames. L'origine de tels biais pourrait donc s'expliquer, dans le cas du déficit en hétérozygotes, par la fixation progressive des loci codant pour les différents caractères précédemment cités et il est possible que RM261 soit plus ou moins lié à un des gènes codant pour ces caractères. C'est ainsi que Ferreira *et al.* (2000) ont constaté, dans la population qu'ils étudiaient (CNA-5), une certaine dérive après 3 cycles de recombinaison sans sélection volontaire et que cette dérive se traduisait, de façon inexpliquée, par un enrichissement progressif en allèles propres à IR36, fournisseur du gène de stérilité mâle. Par ailleurs, les plantes analysées étant, en principe, toutes de génotype ms/+, le gène de stérilité mâle pour lequel la position sur le génome n'est actuellement pas connue, se trouve systématiquement à l'état hétérozygote dans ces plantes. La liaison partielle du locus RM13 ou du locus RM222 avec le gène de stérilité mâle pourrait expliquer l'excès en hétérozygotes constatés à ces loci.

Diversité génétique de la population 00CP

Cette étude visait à étudier la diversité d'une population restauratrice de fertilité utilisée dans un schéma de création de variétés riz hybrides. Elle a permis de resituer cette population vis à vis de la diversité globale des riz cultivés d'origine asiatique, et plus particulièrement des riz *indica*, et vis à vis d'un ensemble de lignées A avec lesquelles elle est amenée à être croisée pour l'obtention d'hybrides. Cette étude est, à notre connaissance, une des premières chez le riz compte tenu du démarrage relativement récent de la sélection récurrente chez cette plante. L'utilisation de marqueurs moléculaires (RAPD et microsatellites) pour assister un programme d'amélioration de population de riz a cependant fait l'objet d'une étude récente (Ferreira *et al.*, 2000). Cette étude a notamment permis d'estimer les distances génétiques entre les lignées fondatrices de la population étudiée et d'apprécier sa diversité génétique. En ce qui concerne notre étude, l'analyse de la diversité allélique de la population 00CP a révélé en moyenne 4,32 allèles par locus et très peu de locus se sont avérés fixés. Cette diversité, en procurant une réserve en allèles relativement

importante, paraît suffisante, au moins dans un premier temps, pour entreprendre une sélection récurrente efficace. La population 00CP a été constituée à l'origine par 36 lignées, mais, l'analyse de leur généalogie, quand on la connaît, montre que des apparentements existent entre ces lignées. Les AFC ont confirmé l'apparentement de la population avec quelques variétés *indica* telles que IR 5, IR 8, IR 36 et Peta. Malheureusement, pour des raisons pratiques, les lignées fondatrices de la population 00CP, à l'exception d'IR36, n'étaient pas disponibles et n'ont pas été intégrées à l'étude. Elles auraient sans doute permis de mieux apprécier la part des différentes origines génétiques des lignées fondatrices et leur niveau de participation dans la population 00CP. La population ayant déjà subi 5 cycles de recombinaison, la comparaison des fréquences alléliques entre le matériel fondateur et la population actuelle aurait permis d'étudier les éventuelles dérives et l'éventuelle apparition d'allèles par mutation ou par migration accidentelle. Par ailleurs, l'origine des allèles spécifiques à la population 00CP n'est pas élucidée. Cela s'explique très probablement par la non-exploration exhaustive de la diversité génétique totale. L'échantillonnage effectué dans chacun des groupes génétiques pour représenter la diversité génétique des riz *O. sativa* est probablement insuffisant pour que l'inventaire exhaustif des allèles ait pu être établi. On ne peut cependant pas écarter l'hypothèse de l'apparition d'allèles par mutation. De part leur nature, les loci microsatellites présentent des fréquences mutationnelles plus élevées que d'autres types de loci. Elles sont estimées entre 10^{-4} et 10^{-5} par locus et par génération alors qu'elles sont estimées à 10^{-6} pour un gène (Jarne et Lagoda, 1996).

Conséquences pour le schéma de création de variétés hybrides

L'inventaire à certains loci des allèles présents dans la population 00CP nous permet dès maintenant de proposer l'introggression de lignées ou variétés dans cette population pour élargir sa base génétique afin d'exploiter au mieux la diversité génétique disponible en adéquation avec les objectifs de la création d'hybrides. Ainsi, l'étude a montré que la population 00CP trouvait son origine génétique à 75 % dans le groupe des variétés *indica* mais qu'elle ne représentait que la moitié de la diversité du groupe *indica*. L'incorporation dans la population de variétés du groupe *indica* présentant des allèles actuellement non présents dans la population devrait permettre à cette population, en la rendant plus représentative de la diversité des riz *indica*, d'accroître à terme son potentiel de progrès vis à vis de son aptitude à la combinaison avec les lignées A. L'apport dans la population 00CP d'allèles spécifiques à d'autres groupes peut aussi être envisagée, notamment en ce qui concerne le groupe des variétés *japonica* pour lequel on retrouve dans la population 15 % d'allèles spécifiques à ce groupe. Du fait de la probable domestication indépendante des groupes *indica* et *japonica* (Second, 1985) et des barrières reproductives qui existent entre ces deux groupes, cette introgression risque cependant d'être difficile (Glaszmann *et al.*, 1999) et de ne pas aboutir aux résultats escomptés. L'utilisation de variétés existantes identifiées comme intermédiaires entre les groupes *indica* et *japonica* ou de populations existantes constituées par un mélange de lignées issues de ces 2 groupes, pourrait probablement faciliter cette introgression. Par ailleurs, l'incorporation de matériel nouveau dans la population devra probablement se faire avec parcimonie et considérée différemment selon l'ampleur de la distance génétique que présentera ce matériel avec la population 00CP. Les variétés incorporées apporteront probablement, outre leur

cortège d'allèles positifs, un ensemble de caractères indésirables (mauvaise architecture de la plante, mauvais format de grain, sensibilité aux maladies,...) réduisant à néant tout le travail de sélection et d'amélioration déjà entrepris.

Par ailleurs, outre la meilleure connaissance de la diversité génétique du matériel étudié, cette étude nous a permis d'estimer les distances génétiques qui existent entre la population restauratrice 00CP et les lignées A, ainsi que celles qui existent entre les lignées A. Un tel résultat devrait permettre, au travers de dispositifs expérimentaux adéquats, de mieux comprendre, à terme, la relation qui existe, chez le riz, entre l'effet d'hétérosis observé chez un hybride et la divergence génétique de ses parents, ainsi que les effets génétiques impliqués dans l'hétérosis. Ainsi, un plan de croisements peut dès maintenant être établi à partir de quelques plantes choisies parmi les 150 plantes de la population 00CP analysées lors de cette étude. En se basant sur les données microsatellites recueillies, ces plantes seront choisies pour représenter au mieux la diversité allélique de la population tout en restreignant au maximum leur nombre. Ce choix pourra, par exemple, être fait selon la méthode de constitution de *core collection* PCS proposée par Noirot *et al.* (1999). Le matériel recombinant issu de ces croisements pourra être à la fois génotypé et évalué pour sa valeur propre et pour sa valeur en croisement avec les lignées A. L'identification de QTL de rendement et des combinaisons alléliques à ces loci conférant un effet d'hétérosis devrait alors être possible. Ainsi, le choix des partenaires restaurateurs sélectionnés à chaque cycle de sélection récurrente pourrait alors être fait en fonction de leur complémentarité allélique avec les lignées A dans le but d'optimiser l'hétérosis.

Remerciements

L'auteur remercie toutes les personnes, et plus particulièrement N. Ahmadi, J. Chantereau, B. Courtois, J.C. Glaszmann, J. Taillebois, ayant contribué à la définition de cette étude, à son bon déroulement et à la mise à disposition des laboratoires et du matériel végétal utilisé. L'auteur remercie également C. Luce pour lui avoir permis d'intégrer ses données à cette étude afin de la rendre plus complète.

Références

- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) GENETIX 4.02, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
- Benzécri, JP (1973). L'Analyse des données. Tome 2 : L'analyse des correspondances. Paris, Dunod, 616 p.
- Bonman JM, Mackill AO, Glaszmann JC (1990) Resistance to *Gerlachia oryzae* in rice. Plant Disease, 74 : 306-309.
- Borrero J, Ospina Y, Guimaraes EP, Chatel M (1997) Ampliación de la base genética de los acervos de arroz, mediante la introducción de variabilidad. In : Selección recurrente en arroz. Guimaraes EP (éd), EMBRAPA, CIRAD-CA, CIAT, Fundación Polar, Cali, Colombia : 55-66.

- Charcosset A, Dubreuil P, Essioux L, Gallais A (1994) Utilisation des techniques de marquage génétique pour l'analyse de la variabilité génétique et la prédiction du phénomène d'hétérosis. *Le sélectionneur français*, 44 : 3-13.
- Charcosset A, Gallais A (1998) Intérêt des marqueurs en sélection. *In* : Les marqueurs moléculaires en génétique et biotechnologies végétales. De Vienne D (éd), INRA : 139-159.
- Chatel M, Guimaraes EP (1995) Seleccion recurrente con androsterilidad en arroz. Cali, Colombie, CIAT, CIAT document n°246, 70 p.
- Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, 95 : 553-567.
- Cheng KS, Wang XK, Zhou JW, Lu YX, Lou J, Huang HW (1984) Studies on indigenous rices in Yunnan and their utilization. 2. A revised classification of Asian cultivated rice. *Acta Agronomica Sinica*, 10 : 271-280.
- Ferreira ME, de O Penteado MI, Brondani C, Belo A, Ferreira MA, N Rangel PH (2000) Caracterizacion y uso de marcadores RAPD y microsatelites (SSR) en el monitoreo del programa de mejoramiento poblacional en arroz. *In* : Avances en el mejoramiento poblacional en arroz. Guimaraes EP (éd), EMBRAPA, Santo Antonio de Goias, Brésil : 37-62.
- Gallais A (1989) Théorie de la sélection en amélioration des plantes. Paris, Masson, 588 p.
- Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. *Theoretical and Applied Genetics*, 74 : 21-30.
- Glaszmann JC, Mew T, Hibino H, Kim CK, Vergel de Dios-Mew TI, Vera Cruz CH, Notteghem JL, Bonman JM (1995) Molecular variation as diverse source of disease resistance in cultivated rice. *In* : Rice Genetics III. Manille, Philippines, IRRI : 460-465.
- Glaszmann JC, Grivet L, Courtois B, Noyer JL, Luce C, Jacquot M, Albar L, Ghesquière A, Second G (1999) Le riz asiatique. *In* : Diversité génétique des plantes tropicales cultivées. Hamon P, Seguin M, Perrier X, Glaszmann JC (éds), Montpellier, France, CIRAD : 43-76.
- Gravois KA, McNew RW (1993) Combining ability and heterosis in U.S. southern long-grain rice. *Crop Science*, 33 : 83-86.
- Guo SW, Thompson EA (1992) Performing the exact test for Hardy-Weinberg proportions for multiple alleles. *Biometrics*, 48 : 361-372.
- Jacquot M, Clément G, Ghesquière A, Glaszmann JC, Guiderdoni E, Tharreau D (1997) Les riz. *In* : L'amélioration des plantes tropicales. Charrier A, Jacquot M, Hamon S, Nicolas D (éds), Montpellier, France, CIRAD : 533-564.
- Jackson MT (1997) Conservation of rice genetic resources : the role of the international rice genebank at IRRI. *Plant Molecular Biology*, 35 : 61-67.
- Jarne P, Lagoda P (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, 11 : 424-429.
- Jones JW (1926) Hybrid vigour in rice. *Journal of American Society of Agronomy*. 18 : 423-428.
- Lin SC, Yuan LP (1980) Hybrid rice breeding in China. *In* : Innovative approaches to rice breeding. IRRI, Los Banos, Laguna, Philippines : 35-51.
- Liu XC, Wu JL (1998) SSR heterogenic patterns of parents for marking and predicting heterosis in rice breeding. *Molecular Breeding*, 4 : 263-268.

- Louis EJ, Dempster ER (1987) An exact test for Hardy-Weinberg and multiple alleles. *Biometrics*, 43 : 805-811.
- Luce C, Noyer JL, Tharreau D, Ahmadi N, Feyt H (2001) The use of microsatellite markers to examine the diversity of the genetic resources of rice (*Oryza sativa*) adapted to european conditions. *In* : Proceedings of the International symposium on molecular markers for characterizing genotypes and identifying cultivars in horticulture. Doré C, Dosba F, Baril C (éds.), Leuven : 221-235.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89 : 583-590.
- Neves PCF, Castro EM, Rangel PHN, Yokoyama LP (1994) Hybrid rice research in Brazil. *In* : Hybrid rice technology : new developments and future prospects. IRRI, Los Banos, Laguna, Philippines : 249-252.
- Noirot M, Anthony F, Dussert S, Hamon S (1999) Une méthode de constitution de *core collections*. *In* : Diversité génétique des plantes tropicales cultivées. Hamon P, Seguin M, Perrier X, Glaszmann JC (éds), Montpellier, France, CIRAD : 77-87.
- Oka HI (1958) Intervarietal variation and classification of cultivated rice. *Indian Journal of Genetics and Plant Breeding*, 18 : 78-89.
- Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSLP) in rice (*Oryza sativa* L.). *Molecular and General Genetics*, 252 : 597-607.
- Patry-Bres C (2000) Diversité génétique et dynamique des formes adventices de riz en Camargue. Thèse de doctorat en sciences, ENSAM, Montpellier, France, 110 p.
- Perrier X, Flori A, Bonnot F (1999) Les méthodes d'analyse de données. *In* : Diversité génétique des plantes tropicales cultivées. Hamon P, Seguin M, Perrier X, Glaszmann JC (éds), Montpellier, France, CIRAD : 43-76.
- Raymond M, Rousset F (1995) GENEPOP (ver. 1.2) : a population genetics software for exact test and ecumenicism. *Journal of Heredity*, 86 : 248-249.
- Risterucci AM, Grivet L, N'Goran JAK, Pieretti I, Flament MH, Lanaud C (2000) A high-density linkage map of *Theobroma cacao* L.. *Theoretical and Applied Genetics*, 101 : 948-955.
- Rousset F, Raymond M (1995) Testing heterozygote excess and deficiency. *Genetics*, 140 : 1413-1419.
- Saghai Maroof MA, Yang GP, Zhang Q, Gravois KA (1997) Correlation between molecular marker distance and hybrid performance in U.S. southern long grain rice. *Crop Science*, 37 : 145-150.
- Second G (1985) Evolutionary relationships in the *sativa* group of *Oryza* based on isozyme data. *Génétique, sélection, évolution*, 17 : 89-114.
- Second G et Ghesquière A (1995) Cartographie des introgressions réciproques entre les sous espèces *indica* et *japonica* de riz cultivé (*Oryza sativa* L.). *In* : Colloque techniques et utilisations des marqueurs moléculaires. Paris, France, Inra, les Colloques de l'Inra n° 72 : 83-93.
- She JX, Autem M, Kotoulas G, Pasteur N, Bonhomme F (1987) Multivariate analysis of genetic exchanges between *Solea aegyptiaca* and *Solea senegalensis* (Teleosts, Soleidae). *Biological Journal of the Linnean Society*, 32 : 357-371.

- Singh RJ, Ikehashi H (1981) Monogenic male-sterility in rice : induction, identification and inheritance. *Crop Science*, 21 : 286-289.
- Sneath PHA, Sokal RR (1973) Numerical taxonomy. San Francisco, Etats-Unis, Freeman, 573 p.
- Taillebois J (1983) Une nouvelle perspective pour la production de semences hybrides F1 : le transfert des caractères d'allogamie de l'espèce *O. longistaminata* A. Chev. à l'espèce *O. sativa* L. *Agronomie Tropicale*, 38 : 303-307.
- Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, 100 : 697-712.
- Virmani SS, Aquino RC, Khush GS (1982) Heterosis breeding in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, 63 : 373-380.
- Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. *Molecular and General Genetics*, 241 : 225-235.
- Xiao J, Li J, Yuan L, Tanksley SD (1995) Dominance is the major genetic basis of heterosis in rice as revealed by QTL analysis using molecular markers. *Genetics*, 140 : 745-754.
- Xiao J, Li J, Yuan L, McCouch SR, Tanksley SD (1996) Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theoretical and Applied Genetics*, 92 : 637-643.
- Yu SB, Li JX, Xu CG, Tan YF, Gao YJ, Li XH, Zhang Q, Saghai Maroof MA (1997) Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences of the USA*, 94 : 9226-9231.
- Zhang Q, Saghai Maroof MA, Lu TY, Shen BZ (1992) Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. *Theoretical and Applied Genetics*, 83 : 495-499.
- Zhang Q, Gao YJ, Yang SH, Ragab RA, Saghai Maroof MA, Li ZB (1994) A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. *Theoretical and Applied Genetics*, 89 : 185-192.
- Zhang QF, Zhou ZQ, Yang GP, Xu CG, Liu KD, Saghai Maroof MA (1996) Molecular marker heterozygosity and hybrid performance in *indica* and *japonica* rice. *Theoretical and Applied Genetics*, 93 : 1218-1224.
- Zhao MF, Li XH, Yang JB, Xu CG, Hu RY, Liu DJ, Zhang Q (1999) Relationship between molecular marker heterozygosity and hybrid performance in intra- and inter-subspecific crosses of rice. *Plant Breeding*, 118 : 139-144.

Annexe : Liste des 85 variétés analysées parmi les 270 variétés de la collection *Oryza sativa* de référence MiniGB

Nom de la variété	N° d'accession IRRI	Origine	Groupe isozymique ^a
AI CHIAO HONG	51250	Chine	<i>indica</i>
AÏCHI ASAHI	40252	Japon	<i>japonica</i>
ARC 13829	42469	Inde	basmati
ASD 1	6267	Inde	<i>indica</i>
ASWINA	26289	Bangladesh	riz flottants précoces
AZUCENA	328	Philippines	<i>japonica</i>
BADKALAMKATI	45011	Inde	<i>indica</i>
BAMOIA 341	6538	Bangladesh	riz flottants précoces
BASMATI 370	6426	Bangladesh	basmati
BHADOIA 233	6541	Bangladesh	riz flottants précoces
BINULAWAN	26872	Philippines	<i>japonica</i>
CAROLINA GOLD	1723	U.S.A.	<i>japonica</i>
CERE AIR	43369	Indonésie	<i>indica</i>
CHAING RONEAL	29726	Cambodge	<i>indica</i>
CHAU	56036	Vietnam	<i>indica</i>
CHIAM CHAM	10214	Vietnam	<i>indica</i>
CICIH BETON	43372	Indonésie	<i>japonica</i>
CO 25	3697	Inde	<i>indica</i>
DA 23 (CHITRAJ)	6208	Bangladesh	<i>indica</i>
DA 9	5854	Bangladesh	<i>indica</i>
DHOLA AMAN	8341	Bangladesh	<i>indica</i>
DOM SOFID	12880	Iran	basmati
DULAR	32561	Inde	aus
FR 13-A	6144	Inde	aus
GOAI	49189	Bangladesh	riz flottants précoces
GOTAK GATIK	43397	Indonésie	<i>japonica</i>
GUAN YIN TSAN	51300	Chine	<i>indica</i>
GWA NGASEIN	33085	Birmanie	<i>indica</i>
HAIFUGOYA	17054	Taiwan	<i>japonica</i>
IL IS AIR	43450	Indonésie	<i>indica</i>
IR5	10321	Philippines	<i>indica</i>
IR8	10320	Philippines	<i>indica</i>
IR36	-	Philippines	<i>indica</i>
IRAT 13	28508	Côte d'Ivoire	<i>japonica</i>
JC 120	9178	Inde	<i>indica</i>
JC 17	9179	Inde	<i>indica</i>
JC 93	9175	Inde	<i>indica</i>
JHONA 349	6307	Inde	aus
JUMULA 2	13375	Népal	<i>japonica</i>
KARKATI 87	6618	Bangladesh	<i>indica</i>
KAUKKYI ANI	33188	Birmanie	basmati
KAW LUYOENG	27716	Thaïlande	<i>indica</i>

Annexe : Suite

Nom de la variété	N° d'accession IRRI	Origine	Groupe isozymique ^a
KHAO DAWK MALI 105	27748	Thaïlande	<i>indica</i>
KIANG CHOU CHIU	8180	Taiwan	<i>indica</i>
LAGEADO	50490	Brésil	<i>indica</i>
LAMBAYQUE 1	10769	Pérou	<i>japonica</i>
LEUANG PRATEW	27762	Thaïlande	<i>indica</i>
MACAN BINUNDOK	8245	Philippines	<i>indica</i>
MAKALIOKA 34	6087	Madagascar	<i>indica</i>
MATIA AMAN 5313	37764	-	riz flottants type Rayada
MOROBEREKAN	12048	Guinée-Conakry	<i>japonica</i>
MTU 9	7919	Inde	<i>indica</i>
N 22	6264	Inde	<i>aus</i>
N 22	19379	Inde	<i>aus</i>
NAM SA GUI 19	11462	Thaïlande	<i>indica</i>
O LUEN CHEUNG	8179	Taiwan	<i>indica</i>
OS 4	11335	Afrique de l'Ouest	<i>japonica</i>
PA TOU HUNG	51400	Chine	<i>indica</i>
PADI HOJONG	8192	Indonésie	<i>indica</i>
PANKHARI 203	5999	Inde	basmati
PAPPAKU	8268	Taiwan	<i>indica</i>
PATIK	43530	Indonésie	<i>indica</i>
PEH KUH TSAO TU	8237	Taiwan	<i>indica</i>
PETA	32571	Indonésie	<i>indica</i>
PIN KAE0	5803	Thaïlande	<i>indica</i>
POPOT 165	43545	Indonésie	<i>indica</i>
PTB 25	6386	Inde	<i>indica</i>
PTB 9	6274	Inde	<i>indica</i>
RATHUWEE	8952	Sri Lanka	<i>indica</i>
RAYADA 16-04	27590	Bangladesh	riz flottants type Rayada
RAYADA 16-05	27591	-	riz flottants type Rayada
RAYADA 16-06	27592	Bangladesh	riz flottants type Rayada
RTS 14	8178	Vietnam	<i>indica</i>
S 624 (AC 398)	8896	Inde	<i>indica</i>
SHAN KIU JU	1154	Chine	<i>japonica</i>
SHORT GRAIN	5075	Thaïlande	<i>indica</i>
SINNA SITHINA KALI	51064	Sri Lanka	<i>indica</i>
SINTANE DIOFOR	5418	Burkina Faso	<i>indica</i>
SOM CAU 70 A	8227	Vietnam	<i>indica</i>
T 26	46768	Inde	basmati
TAM CAU 90 A	8228	Vietnam	<i>indica</i>
TETEP	32576	Vietnam	<i>indica</i>
THIERNO BANDE	16009	Sénégal	<i>indica</i>
TREMBESE	43675	Indonésie	<i>japonica</i>
VARY VATO	10964	Madagascar	<i>indica</i>

^a d'après Glaszmann (1987)